LANS

Fast Forward—Paths to Petaflops

1. What is the rough timetable for feedback of various types for influencing the design?
2. What types of feedback from the consortium would be the most helpful for BG petaflops?
3. How can access be provided to simulators and other tools that will enable the consortium to participate in evaluating BG petaflops design points?
4. What critical tests and experiments/measurements could BG consortium members help with?
5. What systems software issues might be important for BG petaflops timeframe?

What is the rough timetable for feedback of various types for influencing the design?

- Next generation
  - Processor – too late!
  - Network – 3 months
  - Software – longer (3 years ☺)
- Next^2 generation
  - A few years, most things open (but time flies!)

LANS

What types of feedback from the consortium would be the most helpful for BG petaflops?

- Method: Small interest group “brainstorming” in regular meetings with IBM
- IBM has specific questions
  - Consortium can serve as a intermediary
    - Provide high-bandwidth, high-value information exchange with IBM
  - One example: memory traces of applications running on an SMP would be useful for evaluating some design choices
- Performance models of applications
  - Opportunity to bring consortium members together

How can access be provided to simulators and other tools that will enable the consortium to participate in evaluating BG petaflops design points?

- In general, what tools can help in taking the next step beyond “back of the envelope” estimates?
  - Some subsystem simulators such as performance-accurate network simulators may be available
  - Performance-accurate simulators; some consortium members may be able to supply tools

What critical tests and experiments/measurements could BG consortium members help with?

- Quantitative data on the need/value of hybrid MPI+thread (e.g., OpenMP) models
  - Different levels of thread support, such as loop and (the much harder) task
- Scaling constraints of applications
- Quantitative data on the impact of unaligned data at the network
- Failure rates (hardware + software)
  - Consortium partners can collect data
  - Special low-overhead API already exists to simplify problem identification
    - Consortium can encourage developers; software frameworks to include

What systems software issues might be important for BG petaflops timeframe?

- Single System Image
- Fault tolerance; checkpoint/restart
- Programming models
- Load balancing
- Optimizing MPI
  - Consortium activity — Prize for fastest Allreduce ☺